Claims

1. A method for designing a 3-dimentional (3-D) model of a protein, the 3-D representation of at least three family members has already been experimentally obtained, [said 3-D representation presenting similarities], comprising the steps of:

a. identification of common structural blocks (CSBs) among said members of said family,

b. alignment of the amino-acids primary sequence of said family members according to said structural similarities, represented by said CSBs, in order to obtain a first alignment,

c. alignment of said protein as compared on said first alignment, in order to obtain a second alignment, wherein:

i. alignment of said protein is performed in order to optimize the amino-acids alignment between said protein and said first alignment,
when one or more consensus amino-acid exists in said aligned CSBs in said first alignment, and in the amino-acid sequence of said protein, said consensus amino-acids are anchors of said second alignment,

ii. no insertion or deletion of amino-acids can be performed in the aligned CSBs, wherein insertion or deletions are possible in out-of-block regions, if better to align the primary amino-acids sequences,

d. definition of the 3-D structure of CSBs of said protein, according to the 3-D structure of the CSBs of said family members,

e. definition of the global constraints (distance and angular constraints) derived from the comparisons of the structural templates in CSBs, and definition of the local constraints (distance and angular constraints) for the atoms of residues that are not structurally determined after step d. (that are not in the CSBs),

f. selection of rotamers,

g. determination of a family of 3-D model structures of said protein, taking into account said 3-D structure of CSBs obtained in step **d.**, said global and local constraints defined in step **e.**, and said rotamers defined in step **f.**,

h. optimization of said family of 3-D models obtained in step g., by

i discarding structures that present topological defects, and

ii recalculating 3-D structures by taking electrostatic forces into account,

and performing the method again from step **c.** downward, with modifications in the alignment between the primary sequence of said protein and said first alignment, when the obtained model structures do not satisfactorily account for known mutations having biological effects.

2. The method of claim 1, wherein said 3-D representation of family members has been obtained by crystallography or NMR.

3. The method of claim 1, wherein said alignment of said CSBs in step **b.** is performed by use of the GOK software.

4. The method of claim 1, wherein said alignment of said CSBs in step **c.** is performed by use of the GOK software.

5. The method of claim 1, wherein step **d.** is performed according to the following rules:

i. at a given position, when residues are identical between all the template structures and the target sequence, the 3D coordinates of the reference residues are purely assigned to the target residue,

ii. When residues differ, only the coordinates of the backbone atoms are assigned (Cα), and sometimes Cβ or Cγ when they exist.

6. The method of claim 1, wherein said definition of local constraints in step **e.** is performed by analysis of the allowed regions in Ramachandran diagram.

7. The method of claim 1, wherein global and local constraints are selected in step **e.** by the following rules:

i. all distances for which the lower boundary was less than 8 Å.

ii. all the distances involving at least one side-chain atom, to preserve the spatial arrangement between CSBs

iii. all the distances involving atoms of any active group such as an heme group, to fix as much as possible the neighborhood of said active group, such as an iron atom.

8. The method of claim 1, wherein angular constraints are selected in step **e.** by the following rule:

i. dihedral angles ϕ and ψ of all residues located in CSBs are defined as constraints, given by the average values of corresponding ϕ, ψ angles in said family members +/- the calculated standard deviation.

9. The method of claim 1, wherein said rotamers in step **f.** are selected from the couples according to the tables of Dunbrack and Karplus, where the choice of rotamers of a given amino acid is guided by the most favored zones in Ramachandran χ_1, χ_2 maps.

10. The method of claim 1, wherein said step **g.** is performed with the DYANA software.

11. The method of claim 1, wherein said optimization in step **h.** comprises the use of the X-Plor software, as described in A. T. Brunger, X-PLOR, version 3.1.

12. The method of claim 1, wherein said protein is a cytochrome P450 subfamily 3A comprising mammal and human cytochromes P450 3A.

13. The method of claim 12, wherein said mammal cytochrome P450 3A is selected from the group comprising CYP3A6 (SEQ ID N°14), CYP3A12 (SEQ ID N°16), CYP3A29 (SEQ ID N°17) and CYP3A13 (SEQ ID N°18).

14. The method of claim 12, wherein said human cytochrome P450 subfamily 3A is selected from the group comprising CYP3A4 (SEQ ID N°11), CYP3A7 (SEQ ID N°15), CYP3A5 (SEQ ID N°12) and CYP3A43 (SEQ ID N°13).

15. The method of claims 1 and 14, wherein said family members that are used for performing said first alignment for designing a 3-D model of CYP3A4 are chosen from Nor (SEQ ID N° 1), Ery F (SEQ ID N° 2), terp (SEQ ID N° 3), Cam (SEQ ID N° 4), BM3 (SEQ ID N° 5) and 2C5 (SEQ ID N° 6).

16. The method of claims 1 and 14, wherein said family members that are used for performing said first alignment for designing a 3-D model of CYP3A7 are chosen from Ery F (SEQ ID N° 2), BM3 (SEQ ID N° 5), CYP51 (SEQ ID N° 8) and 2C5 (SEQ ID N° 6).

17. A 3-D structure model of a protein, obtained by the method according to claim 1.

18. The model of claim 17, wherein said protein is a cytochrome P450 subfamily 3A comprising mammal and human cytochromes P450 3A.

19. The model of claim 18, wherein said mammal cytochrome P450 3A is selected from the group comprising CYP3A6 (SEQ ID N°14), CYP3A12 (SEQ ID N°16), CYP3A29 (SEQ ID N°17) and CYP3A13 (SEQ ID N°18).

20. The model of claim 18, wherein said human cytochrome P450 subfamily 3A is selected from the group comprising CYP3A4 (SEQ ID N°11), CYP3A7 (SEQ ID N°15), CYP3A5 (SEQ ID N°12) and CYP3A43 (SEQ ID N°13).

21. The model of claim 20, wherein said protein is a cytochrome P450 3A4 or 3A7.

22. The model of claim 21, wherein the residues C97; R104; F101; F107; F247; F303 and C376 are involved in the CYP 3A4 for the recognition and uptake of the substrate at the entry site, and its binding into the active site.

23. The model of claim 20, wherein the residues Q79; F102; R105; R106; F108; F248; F304 and E374 are involved in the CYP 3A7 for the recognition and uptake of the substrate at the entry site, and its binding into the active site.

24. The model of claim 22, having the 3-D atomic coordinates of Table 3.

25 The model of claim 23, having the 3-D atomic coordinates of Table 4.

26. A method for designing a protein, biological functions of which are altered, comprising:

a) obtaining a 3-D model of said protein by the method of claim 1,

b) analyzing said model of step **a.**, and determining the amino-acids that are putatively involved in the biological functions of said protein,

c) changing said amino-acids by mutating the corresponding nucleotides on the nucleic acid sequence coding for said protein, in order to obtain a mutated protein having altered properties.

27. A computer-assisted method for performing restrained dynamics docking of a substrate on an enzyme, a 3-D structure of which is available, comprising the steps of

j. determining a force field, and independently simulating the presence of said enzyme in said force field,

k. minimizing the potential energy (Ep) linked to said force field of said 3-D structure, wherein the spatial position of some atoms of said enzyme is fixed, and wherein the other atoms are mobile, by allowing mobility of the mobile atoms, by

i. simulating an increase in temperature (in order to give kinetic energy),

ii. and minimizing the potential energy by re-specifying the temperature as 0 Kelvin (K)

l. optionally repeating step k in order to obtain other Ep minima, wherein said Ep minima are such that the structure of the protein remains folded,

m. minimizing Ep in said force field of said 3-D structure, wherein all the atoms of the protein are mobile, by

i. simulating an increase in temperature (in order to give kinetic energy), and

ii. minimizing the potential energy by re-specifying the temperature as 0 Kelvin (K)

n. simulating, at 0 K the presence of said substrate next to said enzyme,

o. optionally generating a molecular dynamics simulation on said substrate and enzyme (simulating an increase in temperature, in order to allow mobility of the atoms)

p. generating some constraints to said substrate, in order to impose that it has interaction with said enzyme,

q. generating a molecular dynamics simulation on said substrate and enzyme, with said constraints imposed in step **p.**,

r. optionally, generating a molecular dynamics simulation on said substrate and enzyme without said constraints of step **p.**

28. The method of claim 27, wherein said fixed atoms in step **k.** are the backbone atoms N-Cα-CO in the first minimization step and only Cα in subsequent minimization steps.

29. The method of claim 27, wherein said kinetic energy is simulated by temperature increase to about 100 K for about 5-20 ns.

30. The method of claim 27, wherein said force field in step **j.** comprises forces linked to

a. the distance between atoms,

b. the angles of valence,

c. the dihedral angles,

d. the deformation with regard to planar geometry,

e. the electrostatic field,

f. the Van der Waals forces,

g. hydrogen bonds.

31. The method of claim 27, wherein said constraints in step **p.** are attraction constraints to force said substrate in the active site, and wherein said constraints are not prejudiced to the exact spatial conformation of the substrate in the active site.

32. The method of claim 31, wherein said constraints are final distance constraints between some atoms of said substrate and some atoms of amino-acids present in said active site.

33. The method of claim 27, wherein step **o.** is performed with a simulated temperature of between about 15 and 50 K.

34. The method of claim 27, wherein step **q.** is performed with a simulated temperature of between about 15 and 50 K.

35. The method of claim 27, wherein step **r.** is performed with a simulated temperature of between about 200 and 350 K.

36. The method of claim 27, wherein said enzyme is a cytochrome P450 subfamily 3A comprising mammal and human cytochromes.

37. The method of claim 36, wherein said cytochrome is a cytochrome P450 3A4, and said structure is the structure obtained by the method of claim 15, in particular the model structure of claim 22.

38. The method of claim 36, wherein said substrate is a small organic compound which size can range for example from MW 288 (testosterone) to MW 1203 (cyclosporine A).

39. The method of claim 38, wherein said substrate is testosterone.

40. A computer-assisted method for performing restrained dynamics docking of at least two substrates on an enzyme, a 3-D structure of which is available, comprising the steps consisting of performing the steps depicted in claim 27 with a first substrate and repeating said steps with a second substrate when the first substrate reaches an unconstrained state after molecular dynamics simulation..

41. The method of claim 40, wherein the first and second substrates are the same molecule.

42. The method of claim 40, wherein the first and second substrates are different molecules.

43. The method of claim 41, wherein the first and second substrates display an allosteric effect.

44. The method of claim 41, wherein the first and second substrates display a synergistic effect.

45. The method of claims 41 and 42, wherein at least one of the substrates is an inhibitor or display an inhibitor-based mechanism.

46. The method of claims 41 and 42, wherein at least one of the substrates is an agonist.

47. The method of claim 40 comprising successively repeating the steps of claim 20 with a 3rd, 4th or 5th substrate, some of them being the same or different molecules.

48. The method of claim 40, wherein said fixed atoms in step **k.** are the backbone atoms N-Cα-CO in the first minimization step and only Cα in subsequent minimization steps.

49. The method of claim 40, wherein said kinetic energy is simulated by temperature increase to about 100 K for about 5-20 ns.

50. The method of claim 40, wherein said force field in step **j.** comprises forces linked to:

a. the distance between atoms,

b. the angles of valence

c. the dihedral angles,

d. the deformation with regard to planar geometry,

e. the electrostatic field,

f. the Van der Waals forces

g. hydrogen bonds

51. The method of claim 40, wherein said constraints in step **p.** are attraction constraints to force said substrate in the active site, and wherein said constraints are not prejudiced to the exact spatial conformation of the substrate in the active site.

52. The method of claim 51, wherein said constraints are final distance constraints between some atoms of said substrate and some atoms of amino-acids present in said active site.

53. The method of claim 40, wherein step **o.** is performed with a simulated temperature of between about 15 and 50 K.

54. The method of claim 40, wherein step **q.** is performed with a simulated temperature of between about 15 and 50 K.

55. The method of claim 40, wherein step **r.** is performed with a simulated temperature of between about 200 and 350 K.

56 The method of claim 40, wherein said enzyme is a cytochrome P450 subfamily 3A comprising mammal and human cytochromes P450 3A.

57. The method of claim 56, wherein said cytochrome is cytochrome P450 3A4, and said structure is the structure obtained by the method of claim 15, in particular the model structure of claim 22.

58. The method of claim 40, wherein said first and second substrates are small organic compounds which size can range from MW 288 (testosterone) to MW 1203 (cyclosporine A).

59. The method of claim 58, wherein said substrate is testosterone.

60. The use of the method according to claim 27 or 40 for screening, designing or identifying natural, unnatural substrates or substrate analogs, as well as inhibitors, activators or modulators of said enzyme.

61. The use of the method according to claim 40 or 47 for determining the effect of a first substrate on a second substrate.

62. The use according to claim 61 applied to pharmaceutical products.

63. The use of the method according to claim 40 or 47 for determining the effect of a first testosterone molecule on a second testosterone molecule.

64. The use of the method according to claim 40 or 47 for determining the effect of a first testosterone molecule on a second alpha-naphtoflavone molecule.

65. The use of the method according to claim 27 to 47 for determining the oxidative modification of the substrate according to the proximity to the heme of a part of the substrate.

66. The use of the method according to claim 27 to 39, or 40 to 47, for performing dynamic docking of the said metabolite, either in the absence or in the presence of the second substrate in the computed simulation.

67. The use of the method according to claims 27 to 39, or 40 to 47, to compare the energy of the bound metabolite relatively to the energy of its parent substrate bound, in order to determine if the exit of the given metabolite from the enzyme is favored or not.